UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/753,514	01/04/2001	Yasuyuki Fujikawa	1506.1002 (JDH)	3098

21171 7590 10/27/2006

STAAS & HALSEY LLP
SUITE 700
1201 NEW YORK AVENUE, N.W.
WASHINGTON, DC 20005

EXAMINER	
SAIN, GAUTAM	

ART UNIT	PAPER NUMBER
2176	

DATE MAILED: 10/27/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/753,514	FUJIKAWA, YASUYUKI	
	Examiner	Art Unit	
	Gautam Sain	2176	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *15 August 2006*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,2 and 4-15* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,2 and 4-15* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1) This is a Non-Final rejection in response to the amendments/remarks filed 8/15/2006 (via RCE).

2) Claims 1, 2, 4-15 are pending and rejected in this rejection. Claim 3 was previously cancelled.

3) The effective filing date is 2/4/00 (based on foreign priority).

Continued Examination Under 37 CFR 1.114

4) A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 8/15/2006 has been entered.

Claim Rejections - 35 USC § 103

5) The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5-1) Claims 1, 2 and 4-14 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kuwahara (US 6202072, filed Dec 1997), in view of Nakatsuyama et al (US 5752021, issued May 1998), further in view of Sato et al (US 6014680, issued Jan 11, 2000).

Regarding claim 1, Kuwahara teaches a reading module that reads definition information defining a correlation between elements as basic units configuring the document structure, and defining, for each of the elements (ie., SGML conversion form generation module SGML document read-in module)(col 5, lines 1-18, lines 59-65).

Kuwahara teaches "a retrieving module which refers to the extraction ... target electronic document" (ie., the conversion of a plain text document having a *specific* form to a SGML document; examiner interprets that in order to have a specific form, there must be some condition of specificity in order to perform the conversion which is found out from the document type definition)(col 6, lines 13-36; col 5, lines 50-55 that necessitates a structure for each of the fields – name, department, and address which are the elements – in conjunction with the teachings of Nakatsuyma (see details below)).

Kuwahara teaches "a structure document generating module ... the definition information" (ie., generate a SGML document from a plain text document prepared by a user as part of the two directional conversion between plain text document and a SGML document having a specific form)(col 4, lines 8-13). The examiner interprets the disclosed generating a conversion table for conversion between a unstructured document and a structured document as equivalent to the claimed generating the structured document by adding to each region identifier defined by the definition information because the table serves to correlate fields of the documents using tags (col 2, lines 37-44). The examiner interprets the disclosed tags as equivalent to identified regions.

Kuwahara suggests "wherein said ... lower-order hierarchy", "said retrieving ... higher-order hierarchy" and "said structured document ... lower-order hierarchy" (ie., for conversion from plain text to an SGML document with the document having a specific form by referring to the conversion table, using a document type definition, thereafter adding tags to data that is arranged in a hierarchical manner (where staff encompasses name and department date, as shown in fig 2) and the tags are flank the content data ("development" inputted in the position in between the <Department> and </Department> tags)(col 3, lines 5-24; col 6, lines 2-10; lines 40-50; Fig 2, "Document Type Definition"; Fig 1, items 108 to item 102; Fig 3, item (b) to (c) to (d) shows the plain text, hierarchical structure and SGML output).

Kuwahara in view of Naktsuyama does not expressly teach, but Sato teaches the currently amended limitation of a condition of a pattern within a character string of plain

text data as an extraction condition, and an identifier thereof. For example, Sato discloses a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings (see Sato, Abstract section), where the keyword is a character string from the document expressing a document structure of the non-structured document (col 7, lines 45-50). The examiner interprets the disclosed the extracted keyword as equivalent to the claimed identifier and extraction condition because a keyword serves to extract information using that keyword and a keyword is contained within as character string in the a character string document.

It would have been obvious to one or ordinary skill in the art at the time of the invention to modify Kuwahara to include retrieval conditions on the basis of the retrieval formula for defining the structure of document data as taught by Nakatsuyama, providing the benefit of a document retrieving means to perform retrieval using semantic description and the schema relating to the first schema and directed to the first retrieval and converts the first formula to a second formula (Abstract section), further to include a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a

keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings as taught by Sato, providing the benefit of directly analyzing elements specific to the individual document structure and enabling to directly generate a document instance matching the individual document structure (Sato, col 3, lines 25-31).

Regarding claim 2, Kuwahara teaches a reading module that reads definition information defining a correlation between elements as basic units configuring the document structure, and defining, for each of the elements (ie., SGML conversion form generation module SGML document read-in module)(col 5, lines 1-18, lines 59-65).

Kuwahara teaches "a retrieving module which refers to the extraction ... target electronic document" (ie., the conversion of a plain text document having a *specific* form to a SGML document; examiner interprets that in order to have a specific form, there must be some condition of specificity in order to perform the conversion which is found out from the document type definition)(col 6, lines 13-36; col 5, lines 50-55 that necessitates a structure for each of the fields – name, department, and address which are the elements – in conjunction with the teachings of Nakatsuyma (see details below)).

Kuwahara teaches "a structure document generating module ... the definition information" (ie., generate a SGML document from a plain text document prepared by a user as part of the two directional conversion between plain text document and a SGML document having a specific form)(col 4, lines 8-13). The examiner interprets the

disclosed generating a conversion table for conversion between a unstructured document and a structured document as equivalent to the claimed generating the structured document by adding to each region identifier defined by the definition information because the table serves to correlate fields of the documents using tags (col 2, lines 37-44). The examiner interprets the disclosed tags as equivalent to identified regions.

Kuwahara suggests "said structured document ... lower-order hierarchy" (ie., for conversion from plain text to an SGML document with the document having a specific form by referring to the conversion table, using a document type definition, thereafter adding tags to data that is arranged in a hierarchical manner (where staff encompasses name and department date, as shown in fig 2) and the tags are flank the content data ("development" inputted in the position in between the <Department> and </Department> tags)(col 3, lines 5-24; col 6, lines 2-10; lines 40-50; Fig 2, "Document Type Definition"; Fig 1, items 108 to item 102; Fig 3, item (b) to (c) to (d) shows the plain text, hierarchical structure and SGML output)

Kuwahara in view of Naktsuyama does not expressly teach, but Sato teaches the currently amended limitation of a condition of a pattern within a character string of plain text data as an extraction condition, and an identifier thereof. For example, Sato discloses a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of

the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings (see Sato, Abstract section), where the keyword is a character string from the document expressing a document structure of the non-structured document (col 7, lines 45-50). The examiner interprets the disclosed the extracted keyword as equivalent to the claimed identifier and extraction condition because a keyword serves to extract information using that keyword and a keyword is contained within as character string in the a character string document.

It would have been obvious to one or ordinary skill in the art at the time of the invention to modify Kuwahara to include retrieval conditions on the basis of the retrieval formula for defining the structure of document data as taught by Nakatsuyama; providing the benefit of a document retrieving means to perform retrieval using semantic description and the schema relating to the first schema and directed to the first retrieval and converts the first formula to a second formula (Abstract section), further to include a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings as taught by Sato, providing the benefit of directly analyzing elements specific to the individual document

structure and enabling to directly generate a document instance matching the individual document structure (Sato, col 3, lines 25-31).

Claim 4, Kuwahara suggests "wherein said ... lower-order hierarchy", "said retrieving ... higher-order hierarchy" and "said structured document ... lower-order hierarchy" (ie., for conversion from plain text to an SGML document with the document having a specific form by referring to the conversion table, using a document type definition, thereafter adding tags to data that is arranged in a hierarchical manner (where staff encompasses name and department date, as shown in fig 2) and the tags are flank the content data ("development" inputted in the position in between the <Department> and </Department> tags)(col 3, lines 5-24; col 6, lines 2-10; lines 40-50; Fig 2, "Document Type Definition"; Fig 1, items 108 to item 102; Fig 3, item (b) to (c) to (d) shows the plain text, hierarchical structure and SGML output)

Kuwahara does not expressly teach "repetitive structure" but does suggest it, because Kuwahara teaches that the prior art teaches a repetitive structure (ie., conventional technology generates every time for the entire document type definition processing instead of only once)(col 2, lines 45-50), showing that the repetitive is already well known in the conventional art.

The examiner reasonably interprets the claim limitation of "element in a higher-order hierarchy embraces an element in a lower-order hierarchy that has a repetitive structure" as a structure that has one or more children elements in structure under any given parent element. Based on examiner's interpretations, Kuwhara's disclosure of the document type definition in Fig 2 is equivalent to the claimed limitation because the

document type definition structure of fig 2 shows a parent element "STAFF" (equivalent to higher-order), which embraces two children, "NAME" and "DEPARTMENT" (lower-order). The examiner interprets Kuwahara showing multiple children under one parent as equivalent to a repetitive structure because it is a structure that repeats children elements (lower-order).

Additionally, Kuwahara discloses correlating each of the fields with each of the tags, respectively (col 8, lines 21-31). The examiner interprets this as equivalent to repetitive because the correlations process repeats itself for each of the fields in the document, which implies that it correlates all of the fields and repeats until all of the fields are correlated.

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify Kuwahara to includes generating conversion tables every time (instead of once) as suggest in the prior art of conventional technology, providing the benefit correlating fields of the prototype document to tags of the document type definition when converting plain text to SGML (Kuwahara, col 2, lines 36-44).

Claim 5, Kuwahara suggests "wherein said ... lower-order hierarchy", (ie., for conversion from plain text to an SGML document with the document having a specific form by referring to the conversion table, using a document type definition, thereafter adding tags to data that is arranged in a hierarchical manner (where staff encompasses name and department date, as shown in fig 2) and the tags are flank the content data ("development" inputted in the position in between the <Department> and </Department> tags)(col 3, lines 5-24; col 6, lines 2-10; lines 40-50; Fig 2, "Document

Type Definition"; Fig 1, items 108 to item 102; Fig 3, item (b) to (c) to (d) shows the plain text, hierarchical structure and SGML output).

Kuwahara teaches "said retrieving module extracts [ ] coincident with one higher order hierarchy" (ie., the conversion of a plain text document having a *specific* form to a SGML document; examiner interprets that in order to have a specific form, there must be some condition of specificity in order to perform the conversion which is found out from the document type definition)(col 6, lines 13-36; col 5, lines 50-55 that necessitates a structure for each of the fields – name, department, and address which are the elements – in conjunction with the teachings of Nakatsuyma (see details below))(Fig 2, "Document Type Definition"; Fig 1, items 108 to item 102; Fig 3, item (b) to (c) to (d) shows the plain text, hierarchical structure and SGML output).

Kuwahara does not teach "each region", but does suggest it, because Kuwahara teaches that the prior art teaches a repetitive structure (ie., conventional technology generates every time for the entire document type definition processing instead of only once)(col 2, lines 45-50), showing that the repetitive is already well known in the conventional art and in order for the repetitive processing, different areas have to be processed as separate regions/areas.

Additionally, Kuwahara's Fig 2 illustrates the principles of the operations involved in the preparations of a SGML conversion form based on the plain text document and the document type definition. Specifically, Fig 2 shows a correlation between elements from the plain text document and the document type definition and that the higher order element of "STAFF" embraces lower order elements of "NAME" and "DEPARTMENT".

It would have been obvious to one of ordinary skill in the art at the time of the invention to modify Kuwahara to includes generating conversion tables every time (instead of once) as suggest in the prior art of conventional technology, providing the benefit correlating fields of the prototype document to tags of the document type definition when converting plain text to SGML (Kuwahara, col 2, lines 36-44).

Regarding claims 6 and 7, Kuwahara teaches "extraction condition ... whole region to be extracted" in claim 6 and "extraction condition ... end part thereof" in claim 7 (ie., Plain text document ... "document for Application" and corresponding end tag "document for application")(fig 3, item c)(ie., correlation there between as one unit)(col 5, lines 60-65).

Regarding claims 8 and 9, Kuwahara teaches "description pattern ... to be extracted" (ie., in the plain text document "application form for registering e-mail address"; data displayed)(Fig 3, item a; col 6, lines 23-26).

Regarding claim 10, Kuwahara teaches "extraction condition ... syntax element of the region to be extracted" (ie., text document is analyzed by software for syntax and tags indicating a ... obtained syntax)(col 1, lines 31-40).

Regarding claim 11, Kuwahara teaches a reading module that reads definition information defining a correlation between elements as basic units configuring the document structure, and defining, for each of the elements (ie., SGML conversion form generation module SGML document read-in module)(col 5, lines 1-18, lines 59-65).

Kuwahara teaches "referring to ... reading step", "extracting ... electronic document", "combining the regions ... definition information" (ie., specific form ..

concrete data ... name field, address field; data correlating)(col 6, lines 11-27; col 5, lines 20-30; fig 2; fig 7, item 5).

Kuwahara teaches "generating ... definition information" (col 8, lines 33-38; fig 8). The examiner interprets the disclosed generating a conversion table for conversion between a unstructured document and a structured document as equivalent to the claimed generating the structured document by adding to each region identifier defined by the definition information because the table serves to correlate fields of the documents using tags (col 2, lines 37-44). The examiner interprets the disclosed tags as equivalent to identified regions.

Kuwahara in view of Naktsuyama does not expressly teach, but Sato teaches the currently amended limitation of a condition of a pattern within a character string of plain text data as an extraction condition, and an identifier thereof. For example, Sato discloses a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings (see Sato, Abstract section), where the keyword is a character string from the document expressing a document structure of the non-structured document (col 7, lines 45-50). The examiner interprets the disclosed the extracted keyword as equivalent to the claimed identifier and extraction condition because a keyword serves to extract

information using that keyword and a keyword is contained within as character string in the a character string document.

It would have been obvious to one or ordinary skill in the art at the time of the invention to modify Kuwahara to include retrieval conditions on the basis of the retrieval formula for defining the structure of document data as taught by Nakatsuyama, providing the benefit of a document retrieving means to perform retrieval using semantic description and the schema relating to the first schema and directed to the first retrieval and converts the first formula to a second formula (Abstract section), further to include a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings as taught by Sato, providing the benefit of directly analyzing elements specific to the individual document structure and enabling to directly generate a document instance matching the individual document structure (Sato, col 3, lines 25-31).

Regarding claim 12, Kuwahara teaches a reading module that reads definition information defining a correlation between elements as basic units configuring the document structure, and defining, for each of the elements (ie., SGML conversion form generation module SGML document read-in module)(col 5, lines 1-18, lines 59-65).

Kuwahara teaches *referring to the extraction condition per element that is defined by the reading definition information; extracting a region coincident with the per-element extraction condition referred to out of the processing target electronic document* (ie., prototype file of a plain text document)(col 5, lines 27, fig 2, item 105).

Kuwahara teaches *combining the regions extracted with respect to the respective elements in accordance with the correlation between the respective elements that is defined by the definition information* (ie., generate a SGML document from a plain text document prepared by a user as part of the two directional conversion between plain text document and a SGML document having a specific form)(col 4, lines 8-13).

Kuwahara discloses data correlation of files. Specifically, correlation of a plain text document having a specific form to a SGML document with a document type definition (col 5, lines 25-30; lines 35-40). The Examiner interprets Kuwahara's fields as equivalent to the claimed 'basic unit' (col 6, lines 33-34).

Kuwahara in view of Naktsuyama does not expressly teach, but Sato teaches the currently amended limitation of a condition of a pattern within a character string of plain text data as an extraction condition, and an identifier thereof. For example, Sato discloses a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings (see Sato,

Abstract section), where the keyword is a character string from the document expressing a document structure of the non-structured document (col 7, lines 45-50). The examiner interprets the disclosed the extracted keyword as equivalent to the claimed identifier and extraction condition because a keyword serves to extract information using that keyword and a keyword is contained within as character string in the a character string document.

It would have been obvious to one or ordinary skill in the art at the time of the invention to modify Kuwahara to include retrieval conditions on the basis of the retrieval formula for defining the structure of document data as taught by Nakatsuyama, providing the benefit of a document retrieving means to perform retrieval using semantic description and the schema relating to the first schema and directed to the first retrieval and converts the first formula to a second formula (Abstract section), further to include a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings as taught by Sato, providing the benefit of directly analyzing elements specific to the individual document structure and enabling to directly generate a document instance matching the individual document structure (Sato, col 3, lines 25-31).

Regarding claim 13, Kuwahara teaches *generating the structured document by adding to each region an identifier defined by the definition information* (ie., generate a SGML document from a plain text document prepared by a user as part of the two directional conversion between plain text document and a SGML document having a specific form)(col 4, lines 8-13).

Regarding claim 14, Kuwahara teaches a reading module that reads definition information defining a correlation between elements as basic units configuring the document structure, and defining, for each of the elements (ie., SGML conversion form generation module SGML document read-in module)(col 5, lines 1-18, lines 59-65).

Kuwahara teaches *a retrieving module which refers to the extraction condition per element that is defined by the definition information read by said reading module, and that extracts a region coincident with the per-element extraction condition referred to out of the processing target electronic document* (ie., prototype file of a plain text document)(col 5, lines 27, fig 2, item 105).

Kuwahara in view of Naktsuyama does not expressly teach, but Sato teaches the currently amended limitation of a condition of a pattern within a character string of plain text data as an extraction condition, and an identifier thereof. For example, Sato discloses a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule

which is described by two elements including keywords and other strings (see Sato, Abstract section), where the keyword is a character string from the document expressing a document structure of the non-structured document (col 7, lines 45-50). The examiner interprets the disclosed the extracted keyword as equivalent to the claimed identifier and extraction condition because a keyword serves to extract information using that keyword and a keyword is contained within as character string in the a character string document.

It would have been obvious to one or ordinary skill in the art at the time of the invention to modify Kuwahara to include retrieval conditions on the basis of the retrieval formula for defining the structure of document data as taught by Nakatsuyama, providing the benefit of a document retrieving means to perform retrieval using semantic description and the schema relating to the first schema and directed to the first retrieval and converts the first formula to a second formula (Abstract section), further to include a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings as taught by Sato, providing the benefit of directly analyzing elements specific to the individual document structure and enabling to directly generate a document instance matching the individual document structure (Sato, col 3, lines 25-31).

5-2) Claim 15 is rejected under 35 U.S.C. 103(a) as being unpatentable over Kuwahara (US 6202072, filed Dec 1997), in view of Nakatsuyama et al (US 5752021, issued May 1998).

Regarding claim 15, Kuwahara teaches *a structured document generating module that combines the regions extracted with respect to the respective elements by said retrieving module in accordance with the correlation between the elements that is defined by the definition information* (ie., generate a SGML document from a plain text document prepared by a user as part of the two directional conversion between plain text document and a SGML document having a specific form)(col 4, lines 8-13).

Response to Arguments

Applicant's arguments filed 8/15/06 have been fully considered but they are not persuasive.

Regarding Claims 1-2, 11-12 and 14, The Applicant argues that the combination of references (Kuwahara and Nakatsuyama) does not teach the amended limitation of a condition of a pattern within a character string (Remarks, page 7-8). To deal with this amended limitation, the Examiner introduces the Sato reference. Specifically, Kuwahara in view of Naktsuyama does not expressly teach, but Sato teaches the currently amended limitation of a condition of a pattern within a character string of plain text data as an extraction condition, and an identifier thereof. For example, Sato discloses a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document

structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings (see Sato, Abstract section), where the keyword is a character string from the document expressing a document structure of the non-structured document (col 7, lines 45-50). The examiner interprets the disclosed the extracted keyword as equivalent to the claimed identifier and extraction condition because a keyword serves to extract information using that keyword and a keyword is contained within as character string in the a character string document.

Next, The Applicant argues that Kuwahara does not teach a generation of a structured document by adding to each region an identifier defined by the definition information. The examiner disagrees (Remarks, page 8, middle). Kuwahara discloses the disclosed generating a conversion table for conversion between a unstructured document and a structured document which is equivalent to the claimed generating the structured document by adding to each region identifier defined by the definition information because the table serves to correlate fields of the documents using tags (col 2, lines 37-44). The examiner interprets the disclosed tags as equivalent to identified regions.

Next, Applicant argues that there is no motivation to combine the art in the manner suggested (Remarks, page 8 bottom – page 9, top). The examiner disagrees. To address the claims as amended, it would have been obvious to one or ordinary skill

in the art at the time of the invention to modify Kuwahara to include retrieval conditions on the basis of the retrieval formula for defining the structure of document data as taught by Nakatsuyama, providing the benefit of a document retrieving means to perform retrieval using semantic description and the schema relating to the first schema and directed to the first retrieval and converts the first formula to a second formula (Abstract section), further to include a method for generating a structured document matching the document structure of each non-structured document by using a rule from a present document structure definition for conversion of the non-structured document into the structured document and a keyword extracting module that extracts a keyword representative of the document structure from the non-structured document using a text extraction rule which is described by two elements including keywords and other strings as taught by Sato, providing the benefit of directly analyzing elements specific to the individual document structure and enabling to directly generate a document instance matching the individual document structure (Sato, col 3, lines 25-31). Also, The examiner interprets Kuwahara showing multiple children under one parent as equivalent to a repetitive structure because it is a structure that repeats children elements (lower-order).

Regarding claim 12, The Applicant argues that the combination of references does not define a correlation between elements as basic units (Remarks, page 9). The Examiner disagrees. Kuwahara discloses data correlation of files. Specifically, Kuwahara discloses correlation of a plain text document having a specific form to a SGML document with a document type definition (col 5, lines 25-30; lines 35-40). The

Examiner interprets Kuwahara's fields as equivalent to the claimed 'basic unit' (col 6, lines 33-34) because using the broadest reasonable interpretation a basic unit can be interpreted to include a field.

Regarding Claim 4, The Applicant argues that the combination of references does not teach said correlation between the elements shows a hierarchical structure in which said elements in a higher-order hierarchy embraces an element in a lower-order hierarchy that has a repetitive structure and then retrieving module repeatedly extracts regions (Remarks, page 9-10, section III). The Examiner disagrees. The examiner reasonably interprets the claim limitation of "element in a higher-order hierarchy embraces an element in a lower-order hierarchy that has a repetitive structure" as a structure that has one or more children elements in structure under any given parent element. Based on examiner's interpretations, Kuwhara's disclosure of the document type definition in Fig 2 is equivalent to the claimed limitation because the document type definition structure of fig 2 shows a parent element "STAFF" (equivalent to higher-order), which embraces two children, "NAME" and "DEPARTMENT" (lower-order). The examiner interprets Kuwahara showing multiple children under one parent as equivalent to a repetitive structure because it is a structure that repeats children elements (lower-order).

Additionally, Kuwahara discloses correlating each of the fields with each of the tags, respectively (col 8, lines 21-31). The examiner interprets this as equivalent to repetitive because the correlations process repeats itself for each of the fields in the

document, which implies that it correlates all of the fields and repeats until all of the fields are correlated.

Regarding Claim 5, The Applicant argues that the combination of references does not teach said retrieving module extracts each region coincident with one higher order hierarchy (Remarks, page 10 - page 11).

The Examiner disagrees because Kuwahara's Fig 2 illustrates the principles of the operations involved in the preparations of a SGML conversion form based on the plain text document and the document type definition. Specifically, Fig 2 shows a correlation between elements from the plain text document and the document type definition and that the higher order element of "STAFF" embraces lower order elements of "NAME" and "DEPARTMENT".

Additionally, Regarding Claim 5, The Applicant argues that there is no support for Examiner's assertion for rejection the claim limitation the elements in lower-order hierarchy with reference to the extraction condition of the sequenced element in the lower-order hierarchy out of a region from a portion just after an already-extracted region (see Remarks, page 10) and argues that this limitation would not have been obvious at the time of the invention to modify the art to perform a repetitive process (remarks, pages 10-11). The Examiner disagrees because Kuwahara illustrates the principles of the operations involved in the preparations of a SGML conversion form based on the plain text document and the document type definition. Specifically, Fig 2 shows a correlation between elements from the plain text document and the document type definition and that the higher order element of "STAFF" embraces lower order

elements of "NAME" and "DEPARTMENT". Based on examiner's interpretations, Kuwhara's disclosure of the document type definition in Fig 2 is equivalent to the claimed limitation because the document type definition structure of fig 2 shows a parent element "STAFF" (equivalent to higher-order), which embraces two children, "NAME" and "DEPARTMENT" (lower-order).

Additionally, Kuwahara discloses correlating each of the fields with each of the tags, respectively (col 8, lines 21-31). The examiner interprets this as equivalent to repetitive because the correlations process repeats itself for each of the fields in the document, which implies that it correlates all of the fields and repeats until all of the fields are correlated.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Gautam Sain whose telephone number is 571-272-4096. The examiner can normally be reached on M-F 9-5 EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on 571-272-4136. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

GS 10/23/06
GS

H Herndon
Heather R. Herndon
Supervisory Patent Examiner
Technology Center 2100